

# SEQUENCE LISTING



<110> Genencor International, Inc.  
Bower, Benjamin  
Mitchinson, Colin  
Larenas, Edmund

<120> Cellulase Fusion Protein and Heterologous Cellulase Fusion  
Construct Encoding the Same

<130> GC832-PCT

<140> PCT/US2005/010242

<141> 2005-03-25

<150> US 60/556,711

<151> 2004-03-25

<160> 30

<170> PatentIn version 3.2

<210> 1

<211> 1570

<212> DNA

<213> Trichoderma reesei

<400> 1

atgtatcgga	agttggccgt	catctcggcc	ttcttgccca	cagctcgtgc	tcagtcggcc	60
tgcactctcc	aatcggagac	tcacccgcct	ctgacatggc	agaaatgctc	gtctgggtggc	120
acttgactc	aacagacagg	ctccgtggtc	atcgacgcca	actggcgctg	gactcacgct	180
acgaacagca	gcacgaactg	ctacgatggc	aacacttgga	gctcgaccct	atgtcctgac	240
aacgagacct	gcgcgaagaa	ctgctgtctg	gacgggtgcc	cctacgcgtc	cacgtacgga	300
gttaccacga	gcggtaacag	cctctccatt	ggctttgtca	cccagtctgc	gcagaagaac	360
gttggcgctc	gcctttacct	tatggcgagc	gacacgacct	accaggaatt	caccctgctt	420
ggcaacgagt	tctcttttca	tgttgatggt	tcgcagctgc	cgtaagtgc	ttaccatgaa	480
cccctgacgt	atcttcttgt	gggctcccag	ctgactggcc	aatttaaggt	gcggcttgaa	540
cggagctctc	tacttcgtgt	ccatggacgc	ggatgggtggc	gtgagcaagt	atcccaccaa	600
caccgctggc	gccaaagtacg	gcacggggta	ctgtgacagc	cagtgtcccc	gcgatctgaa	660
gttcatcaat	ggccaggcca	acgttgaggg	ctgggagccg	tcaccaaca	acgcaaacac	720
gggcattgga	ggacacggaa	gctgctgctc	tgagatggat	atctgggagg	ccaactccat	780
ctccgaggct	cttaccctccc	acccttgac	gactgtcggc	caggagatct	gcgaggggtga	840
tgggtgcggc	ggaacttact	ccgataacag	atatggcggc	acttgcgatc	ccgatggctg	900
cgactggaac	ccataccgcc	tgggcaaac	cagcttctac	ggccctggct	caagctttac	960
cctcgatacc	accaagaaat	tgaccgttgt	cacccagttc	gagacgtcgg	gtgccatcaa	1020
ccgatactat	gtccagaatg	gcgtcacttt	ccagcagccc	aacgccgagc	ttggtagtta	1080
ctctggcaac	gagctcaacg	atgattactg	cacagctgag	gaggcagaat	tcggcggatc	1140
ctctttctca	gacaaggggc	gcctgactca	gttcaagaag	gctacctctg	gcggcatggt	1200
tctggctcatg	agtctgtggg	atgatgtgag	tttgatggac	aaacatgcgc	gttgacaaag	1260
agtcaagcag	ctgactgaga	tgttacagta	ctacgccaac	atgctgtggc	tggactccac	1320
ctacccgaca	aacgagacct	cctccacacc	cggtgccgtg	cgcggaagct	gctccaccag	1380
ctccggtgtc	cctgctcagg	tcgaatctca	gtctcccaac	gccaaaggtca	ccttctccaa	1440
catcaagttc	ggaccattg	gcagcaccgg	caaccctagc	ggcggcaacc	ctcccggcgg	1500
aaacccgcct	ggcaccacca	ccacccgccg	cccagccact	accactggaa	gctctcccgg	1560
acctactagt						1570

<210> 2

<211> 51  
 <212> DNA  
 <213> Trichoderma reesei

<400> 2  
 atgtatcggga agttggccgt catctcggcc ttcttggcca cagctcgtgc t 51

<210> 3  
 <211> 1438  
 <212> DNA  
 <213> Trichoderma reesei

<400> 3  
 cagtcggcct gcaactctcca atcggagact caccgcctc tgacatggca gaaatgctcg 60  
 tctggtggca cttgcactca acagacaggc tccgtggtca tcgacgcaa ctggcgctgg 120  
 actcacgcta cgaacagcag cacgaactgc tacgatggca acaactggag ctcgacccta 180  
 tgtcctgaca acgagacctg cgcgaagaac tgctgtctgg acggtgccgc ctacgcgtcc 240  
 acgtacggag ttaccacgag cggtaacagc ctctccattg gctttgtcac ccagtctgcg 300  
 cagaagaacg ttggcgctcg cctttacctt atggcgagcg acacgaccta ccaggaattc 360  
 accctgcttg gcaacgagtt ctctttcgat gttgatgttt cgcagctgcc gtaagtgact 420  
 taccatgaac ccctgacgta tcttcttggtg ggctcccagc tgactggcca atttaagggtg 480  
 cggcttgaac ggagctctct acttcgtgtc catggacgcg gatggtggcg tgagcaagta 540  
 tcccaccaac accgctggcg ccaagtacgg cacggggtac tgtgacagcc agtgtccccg 600  
 cgatctgaag ttcatcaatg gccaggccaa cgttgagggc tgggagccgt catccaacaa 660  
 cgcaaacacg ggcattggag gacacggaag ctgctgctct gagatggata tctgggaggc 720  
 caactccatc tccgaggctc ttacccccca cccttgcaag actgtcggcc aggagatctg 780  
 cgagggtgat ggggtgcggcg gaacttactc cgataacaga tatggcgcca cttgcgatcc 840  
 cgatggctgc gactggaacc cataccgcct gggcaacacc agcttctacg gccctggctc 900  
 aagctttacc ctcgatacca ccaagaaatt gaccgttgtc acccagttcg agacgtcggg 960  
 tgccatcaac cgatactatg tccagaatgg cgtcactttc cagcagccca acgccgagct 1020  
 tggtagttac tctggcaacg agctcaacga tgattactgc acagctgagg aggcagaatt 1080  
 cggcgggatcc tctttctcag acaagggcgg cctgactcag ttcaagaagg ctacctctgg 1140  
 cggcatggtt ctggtcatga gtctgtggga tgatgtgagt ttgatggaca aacatgcgcg 1200  
 ttgacaaaga gtcaagcagc tgactgagat gttacagtac tacgccaaca tgctgtggct 1260  
 ggactccacc taccgcgaaa acgagacctc ctccacacc ggtgccgtgc gcggaagctg 1320  
 ctccaccagc tccggtgtcc ctgctcaggt cgaatctcag tctcccaacg ccaaggtcac 1380  
 cttctccaac atcaagttcg gacccattgg cagcaccggc aaccctagcg gcggcaac 1438

<210> 4  
 <211> 81  
 <212> DNA  
 <213> Trichoderma reesei

<400> 4  
 cctcccggcg gaaacccgcc tggcaccacc accacccgcc gccagccac taccactgga 60  
 agctctcccg gacctactag t 81

<210> 5  
 <211> 480  
 <212> PRT  
 <213> Trichoderma reesei

<400> 5  
 Met Tyr Arg Lys Leu Ala Val Ile Ser Ala Phe Leu Ala Thr Ala Arg  
 1 5 10 15  
 Ala Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr

			20					25					30		
Trp	Gln	Lys	Cys	Ser	Ser	Gly	Gly	Thr	Cys	Thr	Gln	Gln	Thr	Gly	Ser
		35					40					45			
Val	Val	Ile	Asp	Ala	Asn	Trp	Arg	Trp	Thr	His	Ala	Thr	Asn	Ser	Ser
	50					55					60				
Thr	Asn	Cys	Tyr	Asp	Gly	Asn	Thr	Trp	Ser	Ser	Thr	Leu	Cys	Pro	Asp
65					70					75					80
Asn	Glu	Thr	Cys	Ala	Lys	Asn	Cys	Cys	Leu	Asp	Gly	Ala	Ala	Tyr	Ala
				85				90						95	
Ser	Thr	Tyr	Gly	Val	Thr	Thr	Ser	Gly	Asn	Ser	Leu	Ser	Ile	Gly	Phe
			100					105					110		
Val	Thr	Gln	Ser	Ala	Gln	Lys	Asn	Val	Gly	Ala	Arg	Leu	Tyr	Leu	Met
		115					120					125			
Ala	Ser	Asp	Thr	Thr	Tyr	Gln	Glu	Phe	Thr	Leu	Leu	Gly	Asn	Glu	Phe
	130					135					140				
Ser	Phe	Asp	Val	Asp	Val	Ser	Gln	Leu	Pro	Cys	Gly	Leu	Asn	Gly	Ala
145					150					155					160
Leu	Tyr	Phe	Val	Ser	Met	Asp	Ala	Asp	Gly	Gly	Val	Ser	Lys	Tyr	Pro
				165				170						175	
Thr	Asn	Thr	Ala	Gly	Ala	Lys	Tyr	Gly	Thr	Gly	Tyr	Cys	Asp	Ser	Gln
			180					185					190		
Cys	Pro	Arg	Asp	Leu	Lys	Phe	Ile	Asn	Gly	Gln	Ala	Asn	Val	Glu	Gly
		195					200					205			
Trp	Glu	Pro	Ser	Ser	Asn	Asn	Ala	Asn	Thr	Gly	Ile	Gly	Gly	His	Gly
	210				215						220				
Ser	Cys	Cys	Ser	Glu	Met	Asp	Ile	Trp	Glu	Ala	Asn	Ser	Ile	Ser	Glu
225					230					235					240
Ala	Leu	Thr	Pro	His	Pro	Cys	Thr	Thr	Val	Gly	Gln	Glu	Ile	Cys	Glu
				245					250					255	
Gly	Asp	Gly	Cys	Gly	Gly	Thr	Tyr	Ser	Asp	Asn	Arg	Tyr	Gly	Gly	Thr
			260					265					270		
Cys	Asp	Pro	Asp	Gly	Cys	Asp	Trp	Asn	Pro	Tyr	Arg	Leu	Gly	Asn	Thr
		275					280					285			
Ser	Phe	Tyr	Gly	Pro	Gly	Ser	Ser	Phe	Thr	Leu	Asp	Thr	Thr	Lys	Lys
	290					295					300				
Leu	Thr	Val	Val	Thr	Gln	Phe	Glu	Thr	Ser	Gly	Ala	Ile	Asn	Arg	Tyr
305					310					315					320
Tyr	Val	Gln	Asn	Gly	Val	Thr	Phe	Gln	Gln	Pro	Asn	Ala	Glu	Leu	Gly
				325					330					335	
Ser	Tyr	Ser	Gly	Asn	Glu	Leu	Asn	Asp	Asp	Tyr	Cys	Thr	Ala	Glu	Glu
			340					345					350		
Ala	Glu	Phe	Gly	Gly	Ser	Ser	Phe	Ser	Asp	Lys	Gly	Gly	Leu	Thr	Gln
		355					360					365			
Phe	Lys	Lys	Ala	Thr	Ser	Gly	Gly	Met	Val	Leu	Val	Met	Ser	Leu	Trp
	370					375									

<210> 6  
 <211> 431  
 <212> PRT  
 <213> *Trichoderma reesei*

<400> 6  
 Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr Trp  
 1 5 10 15  
 Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser Val  
 20 25 30  
 Val Ile Asp Ala Asn Trp Arg Trp Thr His Ala Thr Asn Ser Ser Thr  
 35 40 45  
 Asn Cys Tyr Asp Gly Asn Thr Trp Ser Ser Thr Leu Cys Pro Asp Asn  
 50 55 60  
 Glu Thr Cys Ala Lys Asn Cys Cys Leu Asp Gly Ala Ala Tyr Ala Ser  
 65 70 75 80  
 Thr Tyr Gly Val Thr Thr Ser Gly Asn Ser Leu Ser Ile Gly Phe Val  
 85 90 95  
 Thr Gln Ser Ala Gln Lys Asn Val Gly Ala Arg Leu Tyr Leu Met Ala  
 100 105 110  
 Ser Asp Thr Thr Tyr Gln Glu Phe Thr Leu Leu Gly Asn Glu Phe Ser  
 115 120 125  
 Phe Asp Val Asp Val Ser Gln Leu Pro Cys Gly Leu Asn Gly Ala Leu  
 130 135 140  
 Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Val Ser Lys Tyr Pro Thr  
 145 150 155 160  
 Asn Thr Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp Ser Gln Cys  
 165 170 175  
 Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Val Glu Gly Trp  
 180 185 190  
 Glu Pro Ser Ser Asn Asn Ala Asn Thr Gly Ile Gly Gly His Gly Ser  
 195 200 205  
 Cys Cys Ser Glu Met Asp Ile Trp Glu Ala Asn Ser Ile Ser Glu Ala  
 210 215 220  
 Leu Thr Pro His Pro Cys Thr Thr Val Gly Gln Glu Ile Cys Glu Gly  
 225 230 235 240  
 Asp Gly Cys Gly Gly Thr Tyr Ser Asp Asn Arg Tyr Gly Gly Thr Cys  
 245 250 255  
 Asp Pro Asp Gly Cys Asp Trp Asn Pro Tyr Arg Leu Gly Asn Thr Ser  
 260 265 270  
 Phe Tyr Gly Pro Gly Ser Ser Phe Thr Leu Asp Thr Thr Lys Lys Leu  
 275 280 285  
 Thr Val Val Thr Gln Phe Glu Thr Ser Gly Ala Ile Asn Arg Tyr Tyr  
 290 295 300  
 Val Gln Asn Gly Val Thr Phe Gln Gln Pro Asn Ala Glu Leu Gly Ser  
 305 310 315 320  
 Tyr Ser Gly Asn Glu Leu Asn Asp Asp Tyr Cys Thr Ala Glu Glu Ala  
 325 330 335  
 Glu Phe Gly Gly Ser Ser Phe Ser Asp Lys Gly Gly Leu Thr Gln Phe  
 340 345 350  
 Lys Lys Ala Thr Ser Gly Gly Met Val Leu Val Met Ser Leu Trp Asp  
 355 360 365  
 Asp Tyr Tyr Ala Asn Met Leu Trp Leu Asp Ser Thr Tyr Pro Thr Asn  
 370 375 380  
 Glu Thr Ser Ser Thr Pro Gly Ala Val Arg Gly Ser Cys Ser Thr Ser  
 385 390 395 400

Ser Gly Val Pro Ala Gln Val Glu Ser Gln Ser Pro Asn Ala Lys Val  
 405 410 415  
 Thr Phe Ser Asn Ile Lys Phe Gly Pro Ile Gly Ser Thr Gly Asn  
 420 425 430

<210> 7  
 <211> 1077  
 <212> DNA  
 <213> Acidothermus cellulolyticus

<400> 7  
 gcgggcggcg gctattggca cagcagcggc cgggagatcc tggacgcgaa caacgtgccg 60  
 gtacggatcg ccggcatcaa ctggttttggg ttcgaaacct gcaattacgt cgtgcacggg 120  
 ctctggtcac gcgactaccg cagcatgctc gaccagataa agtcgctcgg ctacaacaca 180  
 atccggctgc cgtactctga cgacattctc aagccgggca ccatgccgaa cagcatcaat 240  
 ttttaccaga tgaatcagga cctgcagggt ctgacgtcct tgcaggatcat ggacaaaatc 300  
 gtcgcgtacg ccggtcagat cggcctgcgc atcattcttg accgccaccg accggattgc 360  
 agcgggcagt cggcgtctgt gtacacgagc agcgtctcgg aggctacgtg gatttccgac 420  
 ctgcaagcgc tggcgcagcg ctacaaggga aaccgcacgg tcgtcggctt tgacttgac 480  
 aacgagccgc atgaccgcgc ctgctggggc tgcggcgatc cgagcatcga ctggcgattg 540  
 gccgccgagc gggccggaaa cgccgtgctc tcggtgaatc cgaacctgct cattttcgtc 600  
 gaaggtgtgc agagctacaa cggagactcc tactggtggg gcggcaacct gcaaggagcc 660  
 ggccagtacc cggtcgtgct gaacgtgccg aaccgcctgg tgtactcggc gcacgactac 720  
 gcgacgagcg tctaccgcga gacgtggttc agcgatccga ctttcccaa caacatgccc 780  
 ggcatctgga acaagaactg gggatacctc ttcaatcaga acattgcacc ggtatggctg 840  
 ggcgaattcg gtacgacact gcaatccacg accgaccaga cgtggctgaa gacgctcgtc 900  
 cagtacctac ggccgaccgc gcaatacggg gcggacagct tccagtggac cttctgggtc 960  
 tggaaccccg attccggcga cacaggagga attctcaagg atgactggca gacggtcgac 1020  
 acagtaaaag acggctatct cgcgccgac aagtcgtcga ttttcgatcc tgtcggc 1077

<210> 8  
 <211> 359  
 <212> PRT  
 <213> Acidothermus cellulolyticus

<400> 8  
 Ala Gly Gly Gly Tyr Trp His Thr Ser Gly Arg Glu Ile Leu Asp Ala  
 1 5 10 15  
 Asn Asn Val Pro Val Arg Ile Ala Gly Ile Asn Trp Phe Gly Phe Glu  
 20 25 30  
 Thr Cys Asn Tyr Val Val His Gly Leu Trp Ser Arg Asp Tyr Arg Ser  
 35 40 45  
 Met Leu Asp Gln Ile Lys Ser Leu Gly Tyr Asn Thr Ile Arg Leu Pro  
 50 55 60  
 Tyr Ser Asp Asp Ile Leu Lys Pro Gly Thr Met Pro Asn Ser Ile Asn  
 65 70 75 80  
 Phe Tyr Gln Met Asn Gln Asp Leu Gln Gly Leu Thr Ser Leu Gln Val  
 85 90 95  
 Met Asp Lys Ile Val Ala Tyr Ala Gly Gln Ile Gly Leu Arg Ile Ile  
 100 105 110  
 Leu Asp Arg His Arg Pro Asp Cys Ser Gly Gln Ser Ala Leu Trp Tyr  
 115 120 125  
 Thr Ser Ser Val Ser Glu Ala Thr Trp Ile Ser Asp Leu Gln Ala Leu  
 130 135 140  
 Ala Gln Arg Tyr Lys Gly Asn Pro Thr Val Val Gly Phe Asp Leu His  
 145 150 155 160  
 Asn Glu Pro His Asp Pro Ala Cys Trp Gly Cys Gly Asp Pro Ser Ile



gcgctcgcga agacactcga gtactacgcg gcaaaatccg gcgatacggc ctcgcgcgac	1560
ctcgcgaagg gattgctcga ctccatgtgg aacaacgacc aggacagcct cgggtgtgagc	1620
acaccggaga cgcggaccga ctactctcgg ttactcagg tgtacgaccc gacgactggt	1680
gacggcctct acatcccgtc gggttggacg gggaccatgc ccaacggtga ccaaatacaag	1740
ccgggtgcga ccttcctgag catccggtcc tggtacacca aggatccgca gtggtcgaag	1800
gtgcaggcgt acctcaacgg cgggcctgct ccgacgttca actaccaccg gttctgggcg	1860
gagtcgcgact tcgcgatggc gaacgccgat tttggcatgc tcttcccatc cggg	1914

<210> 10

<211> 638

<212> PRT

<213> Acidothermus cellulolyticus

<400> 10

Asn Asp Pro Tyr Ile Gln Arg Phe Leu Thr Met Tyr Asn Lys Ile His	
1 5 10 15	
Asp Pro Ala Asn Gly Tyr Phe Ser Pro Gln Gly Ile Pro Tyr His Ser	
20 25 30	
Val Glu Thr Leu Ile Val Glu Ala Pro Asp Tyr Gly His Glu Thr Thr	
35 40 45	
Ser Glu Ala Tyr Ser Phe Trp Leu Trp Leu Glu Ala Thr Tyr Gly Ala	
50 55 60	
Val Thr Gly Asn Trp Thr Pro Phe Asn Asn Ala Trp Thr Thr Met Glu	
65 70 75 80	
Thr Tyr Met Ile Pro Gln His Ala Asp Gln Pro Asn Asn Ala Ser Tyr	
85 90 95	
Asn Pro Asn Ser Pro Ala Ser Tyr Ala Pro Glu Glu Pro Leu Pro Ser	
100 105 110	
Met Tyr Pro Val Ala Ile Asp Ser Ser Val Pro Val Gly His Asp Pro	
115 120 125	
Leu Ala Ala Glu Leu Gln Ser Thr Tyr Gly Thr Pro Asp Ile Tyr Gly	
130 135 140	
Met His Trp Leu Ala Asp Val Asp Asn Ile Tyr Gly Tyr Gly Asp Ser	
145 150 155 160	
Pro Gly Gly Gly Cys Glu Leu Gly Pro Ser Ala Lys Gly Val Ser Tyr	
165 170 175	
Ile Asn Thr Phe Gln Arg Gly Ser Gln Glu Ser Val Trp Glu Thr Val	
180 185 190	
Thr Gln Pro Thr Cys Asp Asn Gly Lys Tyr Gly Gly Ala His Gly Tyr	
195 200 205	
Val Asp Leu Phe Ile Gln Gly Ser Thr Pro Pro Gln Trp Lys Tyr Thr	
210 215 220	
Asp Ala Pro Asp Ala Asp Ala Arg Ala Val Gln Ala Ala Tyr Trp Ala	
225 230 235 240	
Tyr Thr Trp Ala Ser Ala Gln Gly Lys Ala Ser Ala Ile Ala Pro Thr	
245 250 255	
Ile Ala Lys Ala Ala Lys Leu Gly Asp Tyr Leu Arg Tyr Ser Leu Phe	
260 265 270	
Asp Lys Tyr Phe Lys Gln Val Gly Asn Cys Tyr Pro Ala Ser Ser Cys	
275 280 285	
Pro Gly Ala Thr Gly Arg Gln Ser Glu Thr Tyr Leu Ile Gly Trp Tyr	
290 295 300	
Tyr Ala Trp Gly Gly Ser Ser Gln Gly Trp Ala Trp Arg Ile Gly Asp	
305 310 315 320	
Gly Ala Ala His Phe Gly Tyr Gln Asn Pro Leu Ala Ala Trp Ala Met	
325 330 335	
Ser Asn Val Thr Pro Leu Ile Pro Leu Ser Pro Thr Ala Lys Ser Asp	





```

acgggtgggtc cgtatgacgg gagctccggc gacgtctgga aattctcggg gacctccggg 840
acatggacgc gaatcagccc ggtaccttcg acggacacgg ccaacgacta ctttggttac 900
agcggcctca ctatcgaccg ccagcaccgg aacacgataa tgggtggcaac ccagatatcg 960
tgggtggccgg acaccataat ctttcggagc accgacggcg gtgcgacgtg gacgcggatc 1020
tgggattgga cgagttatcc caatcgaagc ttgcgatatg tgcttgacat ttcggcggag 1080
ccttggtgga ccttcggcgt acagccgaat cctcccgtac cgagtccgaa gctcggctgg 1140
atggatgaag cgatggcaat cgatccgttc aactctgata ggatgctcta cggaacaggc 1200
gcgacgttgt acgcaacaaa tgatctcacg aagtgggact ccggcggcca gattcatatc 1260
gcgcggatgg tcaaaggatt ggaggagacg gcggtaaacg atctcatcag cccgccgtct 1320
ggcgccccgc tcatcagcgc tctcggagac ctcggcggct tcaccacgc cgacgttact 1380
gccgtgccat cgacgatctt cacgtcacgg gtgttcacga ccggcaccag cgtcgactat 1440
gcggaattga atccgtcgat catcgttcgc gctggaagtt tcgatccatc gagccaaccg 1500
aacgacaggc acgtcgcgtt ctcgacagac ggcggcaaga actggttcca aggcagcgaa 1560
cctggcgggg tgacgacggg cggcaccgtc gccgcacgg ccgacggctc tcgtttcgtc 1620
tgggctcccg gcgatcccgg tcagcctgtg gtgtacgcag tcggatttgg caactcctgg 1680
gctgcttcgc aaggtgttcc cgccaatgcc cagatccgct cagaccgggt gaatccaaag 1740
actttctatg ccctatccaa tggaaccttc tatcgaagca cggacggcgg cgtgacattc 1800
caaccggtcg cggccggtct tccgagcagc ggtgccgtcg gtgtcatgtt ccacgcggtg 1860
cctggaaaag aaggcgatct gtggctcgtc gcatcgagcg ggctttacca ctcaaccaat 1920
ggcggcagca gttggtctgc aatcaccggc gtatcctccg cggatgaacgt gggatttggg 1980
aagtctgcgc ccgggtcgtc ataccagcc gtctttgtcg tcggcacgat cggaggcgtt 2040
acgggggctg accgctccga cgacggtggg acgacctggg tacggatcaa tgatgaccag 2100
caccaatacg gaaattgggg acaagcaatc accggtgacc cgcaattta cgggcgggtg 2160
tacataggca cgaacggccg tggaattgtc tacggggaca ttggtggtgc gccgtccgga 2220
tcg

```

<210> 12

<211> 741

<212> PRT

<213> Acidothermus cellulolyticus

<400> 12

```

Ala Thr Thr Gln Pro Tyr Thr Trp Ser Asn Val Ala Ile Gly Gly Gly
1          5          10          15
Gly Phe Val Asp Gly Ile Val Phe Asn Glu Gly Ala Pro Gly Ile Leu
20        25        30
Tyr Val Arg Thr Asp Ile Gly Gly Met Tyr Arg Trp Asp Ala Ala Asn
35        40        45
Gly Arg Trp Ile Pro Leu Leu Asp Trp Val Gly Trp Asn Asn Trp Gly
50        55        60
Tyr Asn Gly Val Val Ser Ile Ala Ala Asp Pro Ile Asn Thr Asn Lys
65        70        75        80
Val Trp Ala Ala Val Gly Met Tyr Thr Asn Ser Trp Asp Pro Asn Asp
85        90        95
Gly Ala Ile Leu Arg Ser Ser Asp Gln Gly Ala Thr Trp Gln Ile Thr
100       105       110
Pro Leu Pro Phe Lys Leu Gly Gly Asn Met Pro Gly Arg Gly Met Gly
115       120       125
Glu Arg Leu Ala Val Asp Pro Asn Asn Asp Asn Ile Leu Tyr Phe Gly
130       135       140
Ala Pro Ser Gly Lys Gly Leu Trp Arg Ser Thr Asp Ser Gly Ala Thr
145       150       155       160
Trp Ser Gln Met Thr Asn Phe Pro Asp Val Gly Thr Tyr Ile Ala Asn
165       170       175
Pro Thr Asp Thr Thr Gly Tyr Gln Ser Asp Ile Gln Gly Val Val Trp
180       185       190
Val Ala Phe Asp Lys Ser Ser Ser Ser Leu Gly Gln Ala Ser Lys Thr

```

		195					200					205				
Ile	Phe	Val	Gly	Val	Ala	Asp	Pro	Asn	Asn	Pro	Val	Phe	Trp	Ser	Arg	
	210					215					220					
Asp	Gly	Gly	Ala	Thr	Trp	Gln	Ala	Val	Pro	Gly	Ala	Pro	Thr	Gly	Phe	
225					230					235					240	
Ile	Pro	His	Lys	Gly	Val	Phe	Asp	Pro	Val	Asn	His	Val	Leu	Tyr	Ile	
				245					250					255		
Ala	Thr	Ser	Asn	Thr	Gly	Gly	Pro	Tyr	Asp	Gly	Ser	Ser	Gly	Asp	Val	
			260				265						270			
Trp	Lys	Phe	Ser	Val	Thr	Ser	Gly	Thr	Trp	Thr	Arg	Ile	Ser	Pro	Val	
		275					280					285				
Pro	Ser	Thr	Asp	Thr	Ala	Asn	Asp	Tyr	Phe	Gly	Tyr	Ser	Gly	Leu	Thr	
	290				295						300					
Ile	Asp	Arg	Gln	His	Pro	Asn	Thr	Ile	Met	Val	Ala	Thr	Gln	Ile	Ser	
305				310						315					320	
Trp	Trp	Pro	Asp	Thr	Ile	Ile	Phe	Arg	Ser	Thr	Asp	Gly	Gly	Ala	Thr	
			325					330						335		
Trp	Thr	Arg	Ile	Trp	Asp	Trp	Thr	Ser	Tyr	Pro	Asn	Arg	Ser	Leu	Arg	
			340				345						350			
Tyr	Val	Leu	Asp	Ile	Ser	Ala	Glu	Pro	Trp	Leu	Thr	Phe	Gly	Val	Gln	
		355					360					365				
Pro	Asn	Pro	Pro	Val	Pro	Ser	Pro	Lys	Leu	Gly	Trp	Met	Asp	Glu	Ala	
	370				375						380					
Met	Ala	Ile	Asp	Pro	Phe	Asn	Ser	Asp	Arg	Met	Leu	Tyr	Gly	Thr	Gly	
385				390						395					400	
Ala	Thr	Leu	Tyr	Ala	Thr	Asn	Asp	Leu	Thr	Lys	Trp	Asp	Ser	Gly	Gly	
			405					410						415		
Gln	Ile	His	Ile	Ala	Pro	Met	Val	Lys	Gly	Leu	Glu	Glu	Thr	Ala	Val	
			420				425						430			
Asn	Asp	Leu	Ile	Ser	Pro	Pro	Ser	Gly	Ala	Pro	Leu	Ile	Ser	Ala	Leu	
		435					440					445				
Gly	Asp	Leu	Gly	Gly	Phe	Thr	His	Ala	Asp	Val	Thr	Ala	Val	Pro	Ser	
	450				455						460					
Thr	Ile	Phe	Thr	Ser	Pro	Val	Phe	Thr	Thr	Gly	Thr	Ser	Val	Asp	Tyr	
465				470						475				480		
Ala	Glu	Leu	Asn	Pro	Ser	Ile	Ile	Val	Arg	Ala	Gly	Ser	Phe	Asp	Pro	
			485					490						495		
Ser	Ser	Gln	Pro	Asn	Asp	Arg	His	Val	Ala	Phe	Ser	Thr	Asp	Gly	Gly	
		500					505						510			
Lys	Asn	Trp	Phe	Gln	Gly	Ser	Glu	Pro	Gly	Gly	Val	Thr	Thr	Gly	Gly	
	515					520						525				
Thr	Val	Ala	Ala	Ser	Ala	Asp	Gly	Ser	Arg	Phe	Val	Trp	Ala	Pro	Gly	
	530				535						540					
Asp	Pro	Gly	Gln	Pro	Val	Val	Tyr	Ala	Val	Gly	Phe	Gly	Asn	Ser	Trp	
545				550												

Val Gly Phe Gly Lys Ser Ala Pro Gly Ser Ser Tyr Pro Ala Val Phe  
660 665 670  
Val Val Gly Thr Ile Gly Gly Val Thr Gly Ala Tyr Arg Ser Asp Asp  
675 680 685  
Gly Gly Thr Thr Trp Val Arg Ile Asn Asp Asp Gln His Gln Tyr Gly  
690 695 700  
Asn Trp Gly Gln Ala Ile Thr Gly Asp Pro Arg Ile Tyr Gly Arg Val  
705 710 715 720  
Tyr Ile Gly Thr Asn Gly Arg Gly Ile Val Tyr Gly Asp Ile Gly Gly  
725 730 735  
Ala Pro Ser Gly Ser  
740

<210> 13  
<211> 1677  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> construct based on *Thermobifida fusca*

<400> 13  
gccggtgct cggaggacta cacggtcaac tcttggggta ccgggttcac cgccaacgtc 60  
accatcacca acctcggcag tgcgatcaac ggctggaccc tggagtggga cttccccggc 120  
aaccagcagg tgaccaacct gtggaacggg acctacaccc agtccgggca gcacgtgtcg 180  
gtcagcaacg ccccgtaaca cgcctccatc ccggccaacg gaacgggtga gttcgggttc 240  
aacggctcct actcggggcag caacgacatc ccctcctcct tcaagctgaa cgggggttacc 300  
tgcgacgggt cggacgaccc cgaccccgag ccagccctc ccccgagccc tccccccagc 360  
cccacagacc cggatgagcc gggcggcccg accaaccgcg ccaccaaccc cggcgagaag 420  
gtcgacaacc cgttcgaggg cgccaagctg tacgtgaacc cggctctggc ggccaaggcc 480  
gccgctgagc cgggcggttc cgcggtcgcc aacgagtcca ccgctgtctg gctggaccgt 540  
atcggcgcca tcgagggcaa cgacagcccg accaccggct ccatgggtct gcgcgaccac 600  
ctggaggagg ccgtccgcca gtccggtggc gacccgctga ccatccaggt cgtcatctac 660  
aacctgcccg gccgcgactg cgccgcgctg gcctccaacg gtgagctggg tcccgatgaa 720  
ctcgaccgct acaagagcga gtacatcgac ccgatcgccg acatcatgtg ggacttcgca 780  
gactacgaga acctgcggat cgtcgccatc atcgagatcg actccctgcc caacctcgtc 840  
accaacgtgg gcgggaacgg cggcaccgag ctctgcgctt acatgaagca gaacggcggc 900  
tacgtcaacg gtgtcggcta cgccctccgc aagctgggag agatcccga cgtctacaac 960  
tacatcgacg ccgcccacca cggctggatc ggctgggact ccaacttcgg cccctcggtg 1020  
gacatcttct acgagggcgc caacgcctcc ggctccaccg tggactacgt gcacggcttc 1080  
atctccaaca cggccaacta ctcgggcact gtggagccgt acctggacgt caacggcacc 1140  
gttaacggcc agctcatccg ccagtccaag tgggttgact ggaaccagta cgtcgacgag 1200  
ctctccttcg tccaggacct gcgtcaggcc ctgatcgcca agggcttcgg gtccgacatc 1260  
ggtatgctca tcgacacctc ccgcaacggc tgggggtggc cgaaccgtcc gaccggaccg 1320  
agtcctcca ccgacctcaa caccacgtt gacgagagcc gtatcgaccg ccgtatccac 1380  
cccggtaact ggtgcaacca ggccggtgag ggccctcgcg agcggcccac ggtcaaccgg 1440  
gctcccgggt ttgacgccta cgtctgggtg aagcccccg gtgagtcgga cggcgccagc 1500  
gaggagatcc cgaacgacga gggcaagggc ttcgaccgca tgtgcgaccc gacctaccag 1560  
ggcaacgccc gcaacggcaa caaccctcg ggtgcgtgc ccaacgcccc catctccggc 1620  
cactgggttct ctgcccagtt ccgcgagctg ctggccaacg cctaccgcgc tctgtaa 1677

<210> 14  
<211> 558  
<212> PRT  
<213> Artificial Sequence

<220>

<223> construct based on Thermobifida fusca

<400> 14

Ala	Gly	Cys	Ser	Val	Asp	Tyr	Thr	Val	Asn	Ser	Trp	Gly	Thr	Gly	Phe
1				5					10					15	
Thr	Ala	Asn	Val	Thr	Ile	Thr	Asn	Leu	Gly	Ser	Ala	Ile	Asn	Gly	Trp
			20					25					30		
Thr	Leu	Glu	Trp	Asp	Phe	Pro	Gly	Asn	Gln	Gln	Val	Thr	Asn	Leu	Trp
		35					40					45			
Asn	Gly	Thr	Tyr	Thr	Gln	Ser	Gly	Gln	His	Val	Ser	Val	Ser	Asn	Ala
	50					55					60				
Pro	Tyr	Asn	Ala	Ser	Ile	Pro	Ala	Asn	Gly	Thr	Val	Glu	Phe	Gly	Phe
65					70					75					80
Asn	Gly	Ser	Tyr	Ser	Gly	Ser	Asn	Asp	Ile	Pro	Ser	Ser	Phe	Lys	Leu
				85				90						95	
Asn	Gly	Val	Thr	Cys	Asp	Gly	Ser	Asp	Asp	Pro	Asp	Pro	Glu	Pro	Ser
			100					105					110		
Pro	Ser	Pro	Ser	Pro	Ser	Pro	Ser	Pro	Thr	Asp	Pro	Asp	Glu	Pro	Gly
		115					120					125			
Gly	Pro	Thr	Asn	Pro	Pro	Thr	Asn	Pro	Gly	Glu	Lys	Val	Asp	Asn	Pro
	130					135						140			
Phe	Glu	Gly	Ala	Lys	Leu	Tyr	Val	Asn	Pro	Val	Trp	Ser	Ala	Lys	Ala
145					150					155					160
Ala	Ala	Glu	Pro	Gly	Gly	Ser	Ala	Val	Ala	Asn	Glu	Ser	Thr	Ala	Val
				165					170					175	
Trp	Leu	Asp	Arg	Ile	Gly	Ala	Ile	Glu	Gly	Asn	Asp	Ser	Pro	Thr	Thr
			180					185					190		
Gly	Ser	Met	Gly	Leu	Arg	Asp	His	Leu	Glu	Glu	Ala	Val	Arg	Gln	Ser
		195					200					205			
Gly	Gly	Asp	Pro	Leu	Thr	Ile	Gln	Val	Val	Ile	Tyr	Asn	Leu	Pro	Gly
	210					215					220				
Arg	Asp	Cys	Ala	Ala	Leu	Ala	Ser	Asn	Gly	Glu	Leu	Gly	Pro	Asp	Glu
225					230					235					240
Leu	Asp	Arg	Tyr	Lys	Ser	Glu	Tyr	Ile	Asp	Pro	Ile	Ala	Asp	Ile	Met
				245					250					255	
Trp	Asp	Phe	Ala	Asp	Tyr	Glu	Asn	Leu	Arg	Ile	Val	Ala	Ile	Ile	Glu
			260					265					270		
Ile	Asp	Ser	Leu	Pro	Asn	Leu	Val	Thr	Asn	Val	Gly	Gly	Asn	Gly	Gly
		275					280					285			
Thr	Glu	Leu	Cys	Ala	Tyr	Met	Lys	Gln	Asn	Gly	Gly	Tyr	Val	Asn	Gly
	290					295					300				
Val	Gly	Tyr	Ala	Leu	Arg	Lys	Leu	Gly	Glu	Ile	Pro	Asn	Val	Tyr	Asn
305					310					315					320
Tyr	Ile	Asp	Ala	Ala	His	His	Gly	Trp	Ile	Gly	Trp	Asp	Ser	Asn	Phe
				325					330					335	
Gly	Pro	Ser	Val	Asp	Ile	Phe	Tyr	Glu	Ala	Ala	Asn	Ala	Ser	Gly	Ser
			340					345					350		
Thr	Val	Asp	Tyr	Val	His	Gly	Phe	Ile	Ser	Asn	Thr	Ala	Asn	Tyr	Ser
		355					360					365			
Ala	Thr	Val	Glu	Pro	Tyr	Leu	Asp	Val	Asn	Gly	Thr	Val	Asn	Gly	Gln
	370					375					380				
Leu	Ile	Arg	Gln	Ser	Lys	Trp	Val	Asp	Trp	Asn	Gln	Tyr	Val	Asp	Glu
385					390					395					400
Leu	Ser	Phe	Val	Gln	Asp	Leu	Arg	Gln	Ala	Leu	Ile	Ala	Lys	Gly	Phe
				405					410					415	
Arg	Ser	Asp	Ile	Gly	Met	Leu	Ile	Asp	Thr	Ser	Arg	Asn	Gly	Trp	Gly
			420					425					430		

Gly	Pro	Asn	Arg	Pro	Thr	Gly	Pro	Ser	Ser	Ser	Thr	Asp	Leu	Asn	Thr
		435					440					445			
Tyr	Val	Asp	Glu	Ser	Arg	Ile	Asp	Arg	Arg	Ile	His	Pro	Gly	Asn	Trp
	450					455					460				
Cys	Asn	Gln	Ala	Gly	Ala	Gly	Leu	Gly	Glu	Arg	Pro	Thr	Val	Asn	Pro
465					470					475					480
Ala	Pro	Gly	Val	Asp	Ala	Tyr	Val	Trp	Val	Lys	Pro	Pro	Gly	Glu	Ser
			485					490						495	
Asp	Gly	Ala	Ser	Glu	Glu	Ile	Pro	Asn	Asp	Glu	Gly	Lys	Gly	Phe	Asp
			500					505					510		
Arg	Met	Cys	Asp	Pro	Thr	Tyr	Gln	Gly	Asn	Ala	Arg	Asn	Gly	Asn	Asn
		515					520					525			
Pro	Ser	Gly	Ala	Leu	Pro	Asn	Ala	Pro	Ile	Ser	Gly	His	Trp	Phe	Ser
	530					535					540				
Ala	Gln	Phe	Arg	Glu	Leu	Leu	Ala	Asn	Ala	Tyr	Pro	Pro	Leu		
545					550					555					

<210> 15  
 <211> 1293  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> construct based on *Thermobifida fusca*

<400> 15

gcccgtctca	ccgccacagt	caccaaagaa	tcctcgtggg	acaacggcta	ctccgcgtcc	60
gtcaccgtcc	gcaacgacac	ctcgagcacc	gtctcccagt	gggaggtcgt	cctcaccctg	120
cccggcggga	ctacagtggc	ccaggtgtgg	aacgcccagc	acaccagcag	cggcaactcc	180
cacaccttca	ccgggggttc	ctggaacagc	accatcccgc	ccggaggcac	cgcctcttcc	240
ggcttcatcg	cttccaggcag	cggcgaaccc	acccactgca	ccatcaacgg	cgccccctgc	300
gacgaaggct	ccgagccggg	cggccccggc	gggtcccggaa	ccccctcccc	cgaccccggc	360
acgcagcccc	gcaccggcac	cccggtcgag	cggtagcgca	aagtccaggt	ctgcggcacc	420
cagctctgcg	acgagcacgg	caaccgggtc	caactgcgcg	gcatgagcac	ccacggcatc	480
cagtggttcg	accactgcct	gaccgacagc	tcgctggacg	ccctggccta	cgactggaag	540
gccgacatca	tcgcctgtc	catgtacatc	caggaagacg	gctacgagac	caaccgcgcg	600
ggcttcaccg	accggatgca	ccagctcatc	gacatggcca	cggcgcgcgg	cctgtacgtg	660
atcgtggact	ggcacatcct	caccccgggc	gatccccact	acaacctgga	ccggggccaag	720
accttcttcg	cggaaatcgc	ccagcgccac	gccagcaaga	ccaacgtgct	ctacgagatc	780
gccaacgaac	ccaacggagt	gagctggggc	tccatcaaga	gctacgccga	agaggtcatc	840
ccggtgatcc	gccagcgcca	ccccgactcg	gtgatcatcg	tgggcacccg	cggctggctg	900
tcgctcggcg	tctccgaagg	ctccggcccc	gccgagatcg	cggccaaccc	ggtcaacgcc	960
tccaacatca	tgtacgcctt	ccacttctac	gcggcctcgc	accgcgacaa	ctacctcaac	1020
gcgctgcgtg	aggcctccga	gctgttcccg	gtcttcgtca	ccgagttcgg	caccgagacc	1080
tacaccggtg	acggcgccaa	cgacttcag	atggccgacc	gctacatcga	cctgatggcg	1140
gaacggaaga	tcgggtggac	caagtggaac	tactcggacg	acttcggttc	cggcgcggtc	1200
ttccagccgg	gcacctgcgc	gtccggcggc	ccgtggagcg	gttcgtcgct	gaaggcgtcc	1260
ggacagtggg	tgcggagcaa	gctccagtcc	tga			1293

<210> 16  
 <211> 430  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> construct based on *Thermobifida fusca*

<400> 16

Ala	Gly	Leu	Thr	Ala	Thr	Val	Thr	Lys	Glu	Ser	Ser	Trp	Asp	Asn	Gly		
1				5					10					15			
Tyr	Ser	Ala	Ser	Val	Thr	Val	Arg	Asn	Asp	Thr	Ser	Ser	Thr	Val	Ser		
			20					25					30				
Gln	Trp	Glu	Val	Val	Leu	Thr	Leu	Pro	Gly	Gly	Thr	Thr	Val	Ala	Gln		
		35					40					45					
Val	Trp	Asn	Ala	Gln	His	Thr	Ser	Ser	Gly	Asn	Ser	His	Thr	Phe	Thr		
	50					55					60						
Gly	Val	Ser	Trp	Asn	Ser	Thr	Ile	Pro	Pro	Gly	Gly	Thr	Ala	Ser	Ser		
65					70					75					80		
Gly	Phe	Ile	Ala	Ser	Gly	Ser	Gly	Glu	Pro	Thr	His	Cys	Thr	Ile	Asn		
				85				90						95			
Gly	Ala	Pro	Cys	Asp	Glu	Gly	Ser	Glu	Pro	Gly	Gly	Pro	Gly	Gly	Pro		
			100					105					110				
Gly	Thr	Pro	Ser	Pro	Asp	Pro	Gly	Thr	Gln	Pro	Gly	Thr	Gly	Thr	Pro		
		115					120					125					
Val	Glu	Arg	Tyr	Gly	Lys	Val	Gln	Val	Cys	Gly	Thr	Gln	Leu	Cys	Asp		
	130					135					140						
Glu	His	Gly	Asn	Pro	Val	Gln	Leu	Arg	Gly	Met	Ser	Thr	His	Gly	Ile		
145					150					155					160		
Gln	Trp	Phe	Asp	His	Cys	Leu	Thr	Asp	Ser	Ser	Leu	Asp	Ala	Leu	Ala		
				165				170						175			
Tyr	Asp	Trp	Lys	Ala	Asp	Ile	Ile	Arg	Leu	Ser	Met	Tyr	Ile	Gln	Glu		
			180					185					190				
Asp	Gly	Tyr	Glu	Thr	Asn	Pro	Arg	Gly	Phe	Thr	Asp	Arg	Met	His	Gln		
		195					200					205					
Leu	Ile	Asp	Met	Ala	Thr	Ala	Arg	Gly	Leu	Tyr	Val	Ile	Val	Asp	Trp		
	210					215					220						
His	Ile	Leu	Thr	Pro	Gly	Asp	Pro	His	Tyr	Asn	Leu	Asp	Arg	Ala	Lys		
225					230					235					240		
Thr	Phe	Phe	Ala	Glu	Ile	Ala	Gln	Arg	His	Ala	Ser	Lys	Thr	Asn	Val		
			245					250						255			
Leu	Tyr	Glu	Ile	Ala	Asn	Glu	Pro	Asn	Gly	Val	Ser	Trp	Ala	Ser	Ile		
		260						265					270				
Lys	Ser	Tyr	Ala	Glu	Glu	Val	Ile	Pro	Val	Ile	Arg	Gln	Arg	Asp	Pro		
		275					280					285					
Asp	Ser	Val	Ile	Ile	Val	Gly	Thr	Arg	Gly	Trp	Ser	Ser	Leu	Gly	Val		
	290					295					300						
Ser	Glu	Gly	Ser	Gly	Pro	Ala	Glu	Ile	Ala	Ala	Asn	Pro	Val	Asn	Ala		
305					310					315					320		
Ser	Asn	Ile	Met	Tyr	Ala	Phe	His	Phe	Tyr	Ala	Ala	Ser	His	Arg	Asp		
			325					330						335			
Asn	Tyr	Leu	Asn	Ala	Leu	Arg	Glu	Ala	Ser	Glu	Leu	Phe	Pro	Val	Phe		
		340						345					350				
Val	Thr	Glu	Phe	Gly	Thr	Glu	Thr	Tyr	Thr	Gly	Asp	Gly	Ala	Asn	Asp		
		355					360					365					
Phe	Gln	Met	Ala	Asp	Arg	Tyr	Ile	Asp	Leu	Met	Ala	Glu	Arg	Lys	Ile		
	370					375					380						
Gly	Trp	Thr	Lys	Trp	Asn	Tyr	Ser	Asp	Asp	Phe	Arg	Ser	Gly	Ala	Val		
385					390					395					400		
Phe	Gln	Pro	Gly	Thr	Cys	Ala	Ser	Gly	Gly	Pro	Trp	Ser	Gly	Ser	Ser		
			405					410						415			
Leu	Lys	Ala	Ser	Gly	Gln	Trp	Val	Arg	Ser	Lys	Leu	Gln	Ser				
		420						425					430				

<210> 17

<211> 2656  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> fusion construct

<400> 17  
 atgtatcggga agttggccgt catctcggcc ttcttggcca cagctcgtgc tcagtcggcc 60  
 tgcactctcc aatcggagac tcacccgcct ctgacatggc agaaatgctc gtctgggtggc 120  
 acttgcactc aacagacagg ctccgtggtc atcgacgcca actggcgctg gactcacgct 180  
 acgaacagca gcacgaactg ctacgatggc aacacttgga gctcgaccct atgtcctgac 240  
 aacgagacct gcgcgaagaa ctgctgtctg gacgggtgccc cctacgcgtc cacgtacgga 300  
 gttaccacga gcggttaacag cctctccatt ggctttgtca cccagtctgc gcagaagaac 360  
 gttggcgctc gcctttacct tatggcgagc gacacgacct accaggaatt caccctgctt 420  
 ggcaacgagt tctctttcga tgttgatgtt tcgcagctgc cgtaagtgc ttaccatgaa 480  
 cccctgacgt atcttcttgt gggctcccag ctgactggcc aatttaaggt gcggcttgaa 540  
 cggagctctc tacttctgtt ccatggacgc ggatgggtggc gtgagcaagt atcccaccaa 600  
 caccgctggc gccaaagtacg gcacggggta ctgtgacagc cagtgtcccc gcgatctgaa 660  
 gttcatcaat ggccaggcca acgttgaggg ctgggagccg tcatccaaca acgcaaacac 720  
 gggcatttga ggacacggaa gctgctgctc tgagatggat atctgggagg ccaactccat 780  
 ctccgaggct cttaccccc acccttgca gactgtcggc caggagatct gcgaggggtga 840  
 tgggtgcggc ggaacttact ccgataacag atatggcggc acttgcgatc ccgatggctg 900  
 cgactggaac ccataccgcc tgggcaacac cagcttctac ggccctggct caagctttac 960  
 cctcgatacc accaagaaat tgaccgttgt caccagttc gagacgtcgg gtgccatcaa 1020  
 ccgatactat gtccagaatg gcgtcacttt ccagcagccc aacgccgagc ttggtagtta 1080  
 ctcttgcaac gagctcaacg atgattactg cacagctgag gaggcagaat tcggcggatc 1140  
 ctctttctca gacaaggggc gcctgactca gttcaagaag gctacctctg gcggcattgt 1200  
 tctgggtcatg agtctgtggg atgatgtgag tttgatggac aaacatgcgc gttgacaaag 1260  
 agtcaagcag ctgactgaga tgttacagta ctacgccaac atgctgtggc tggactccac 1320  
 ctacccgaca aacgagacct cctccacacc cgggtgccgtg cgcggaagct gctccaccag 1380  
 ctccggtgtc cctgtcaggg tcgaatctca gtctcccaac gccaaaggtca ctttctccaa 1440  
 catcaagttc ggacccattg gcagcaccgg caaccctagc ggcggaacc ctcccgcg 1500  
 aaaccgcctt ggcaccacca ccaccgcgg cccagccact accactggaa gctctcccg 1560  
 acctactagt aagcggggcg gcggcggtta ttggcacac agcggccggg agatcctgga 1620  
 cgcgaacaac gtgcccgtac ggatcgccgg catcaactgg tttgggttcg aaacctgca 1680  
 ttacgtcgtg cacggtctct ggtcacgcga ctaccgcagc atgctcgacc agataaagtc 1740  
 gctcggctac aacacaatcc gggtgcccga ctctgacgac attctcaagc cgggcaccat 1800  
 gccgaacagc atcaattttt accagatgaa tcaggacctg caggggtctga cgtccttgca 1860  
 ggtcatggac aaaatcgtcg cgtacgccgg tcagatcggc ctgcgcatca ttcttgaccg 1920  
 ccaccgaccg gattgcagcg ggcagtcggc gctgtggtac acgagcagcg tctcggaggc 1980  
 tacgtggatt tccgacctgc aagcgtggc gcagcgctac aagggaaacc cgacggtcgt 2040  
 cggttttgac ttgcacaacg agccgcatga cccggcctgc tggggctgcg gcgatccgag 2100  
 catcgactgg cgattggccg ccgagcgggc cggaaacgcc gtgctctcgg tgaatccgaa 2160  
 cctgctcatt ttcgtcgaag gtgtgcagag ctacaacgga gactcctact ggtggggcg 2220  
 caacctgcaa ggagccggcc agtaccgggt cgtgctgaac gtgccgaacc gcctggtgta 2280  
 ctcggcgcac gactacgca cgagcgtcta cccgcagacg tggttcagcg atccgacctt 2340  
 ccccaacaac atgcccggca tctggaacaa gaactgggga tacctcttca atcagaacat 2400  
 tgcaccggta tggctggcg aattcgttac gacactgcaa tccacgaccg accagactg 2460  
 gctgaagacg ctcgctccagt acctacggcc gaccgcgcaa tacggtgagg acagcttcca 2520  
 gtggaccttc tggctcctgga accccgattc cggcgacaca ggaggaattc tcaaggatga 2580  
 ctggcagacg gtcgacacag taaaagacgg ctatctcgcg ccgatcaagt cgtcgatttt 2640  
 cgatcctgtc ggctaa 2656

<210> 18  
 <211> 841  
 <212> PRT

<213> Artificial Sequence

<220>

<223> fusion construct

<400> 18

Met	Tyr	Arg	Lys	Leu	Ala	Val	Ile	Ser	Ala	Phe	Leu	Ala	Thr	Ala	Arg
1			5						10					15	
Ala	Gln	Ser	Ala	Cys	Thr	Leu	Gln	Ser	Glu	Thr	His	Pro	Pro	Leu	Thr
			20					25				30			
Trp	Gln	Lys	Cys	Ser	Ser	Gly	Gly	Thr	Cys	Thr	Gln	Gln	Thr	Gly	Ser
		35					40					45			
Val	Val	Ile	Asp	Ala	Asn	Trp	Arg	Trp	Thr	His	Ala	Thr	Asn	Ser	Ser
	50					55					60				
Thr	Asn	Cys	Tyr	Asp	Gly	Asn	Thr	Trp	Ser	Ser	Thr	Leu	Cys	Pro	Asp
65					70					75					80
Asn	Glu	Thr	Cys	Ala	Lys	Asn	Cys	Cys	Leu	Asp	Gly	Ala	Ala	Tyr	Ala
				85					90					95	
Ser	Thr	Tyr	Gly	Val	Thr	Thr	Ser	Gly	Asn	Ser	Leu	Ser	Ile	Gly	Phe
			100					105					110		
Val	Thr	Gln	Ser	Ala	Gln	Lys	Asn	Val	Gly	Ala	Arg	Leu	Tyr	Leu	Met
			115				120					125			
Ala	Ser	Asp	Thr	Thr	Tyr	Gln	Glu	Phe	Thr	Leu	Leu	Gly	Asn	Glu	Phe
	130					135					140				
Ser	Phe	Asp	Val	Asp	Val	Ser	Gln	Leu	Pro	Cys	Gly	Leu	Asn	Gly	Ala
145					150					155					160
Leu	Tyr	Phe	Val	Ser	Met	Asp	Ala	Asp	Gly	Gly	Val	Ser	Lys	Tyr	Pro
			165						170					175	
Thr	Asn	Thr	Ala	Gly	Ala	Lys	Tyr	Gly	Thr	Gly	Tyr	Cys	Asp	Ser	Gln
			180					185					190		
Cys	Pro	Arg	Asp	Leu	Lys	Phe	Ile	Asn	Gly	Gln	Ala	Asn	Val	Glu	Gly
		195					200					205			
Trp	Glu	Pro	Ser	Ser	Asn	Asn	Ala	Asn	Thr	Gly	Ile	Gly	Gly	His	Gly
	210					215					220				
Ser	Cys	Cys	Ser	Glu	Met	Asp	Ile	Trp	Glu	Ala	Asn	Ser	Ile	Ser	Glu
225					230					235					240
Ala	Leu	Thr	Pro	His	Pro	Cys	Thr	Thr	Val	Gly	Gln	Glu	Ile	Cys	Glu
				245					250					255	
Gly	Asp	Gly	Cys	Gly	Gly	Thr	Tyr	Ser	Asp	Asn	Arg	Tyr	Gly	Gly	Thr
			260					265					270		
Cys	Asp	Pro	Asp	Gly	Cys	Asp	Trp	Asn	Pro	Tyr	Arg	Leu	Gly	Asn	Thr
	275						280					285			
Ser	Phe	Tyr	Gly	Pro	Gly	Ser	Ser	Phe	Thr	Leu	Asp	Thr	Thr	Lys	Lys
	290					295					300				
Leu	Thr	Val	Val	Thr	Gln	Phe	Glu	Thr	Ser	Gly	Ala	Ile	Asn	Arg	Tyr
305					310					315					320
Tyr	Val	Gln	Asn	Gly	Val	Thr	Phe	Gln	Gln	Pro	Asn	Ala	Glu	Leu	Gly
			325						330					335	
Ser	Tyr	Ser	Gly	Asn	Glu	Leu	Asn	Asp	Asp	Tyr	Cys	Thr	Ala	Glu	Glu
			340					345					350		
Ala	Glu	Phe	Gly	Gly	Ser	Ser	Phe	Ser	Asp	Lys	Gly	Gly	Leu	Thr	Gln
		355					360					365			
Phe	Lys	Lys	Ala	Thr	Ser	Gly	Gly	Met	Val	Leu	Val	Met	Ser	Leu	Trp
	370					375					380				
Asp	Asp	Tyr	Tyr	Ala	Asn	Met	Leu	Trp	Leu	Asp	Ser	Thr	Tyr	Pro	Thr
385					390					395					400
Asn	Glu	Thr	Ser	Ser	Thr	Pro	Gly	Ala	Val	Arg	Gly	Ser	Cys	Ser	Thr



				405					410					415			
Ser	Ser	Gly	Val	Pro	Ala	Gln	Val	Glu	Ser	Gln	Ser	Pro	Asn	Ala	Lys		
			420					425					430				
Val	Thr	Phe	Ser	Asn	Ile	Lys	Phe	Gly	Pro	Ile	Gly	Ser	Thr	Gly	Asn		
		435					440					445					
Pro	Ser	Gly	Gly	Asn	Pro	Pro	Gly	Gly	Asn	Pro	Pro	Gly	Thr	Thr	Thr		
	450					455					460						
Thr	Arg	Arg	Pro	Ala	Thr	Thr	Thr	Gly	Ser	Ser	Pro	Gly	Pro	Thr	Ser		
465					470					475					480		
Lys	Arg	Ala	Gly	Gly	Gly	Tyr	Trp	His	Thr	Ser	Gly	Arg	Glu	Ile	Leu		
			485					490					495				
Asp	Ala	Asn	Asn	Val	Pro	Val	Arg	Ile	Ala	Gly	Ile	Asn	Trp	Phe	Gly		
		500						505				510					
Phe	Glu	Thr	Cys	Asn	Tyr	Val	Val	His	Gly	Leu	Trp	Ser	Arg	Asp	Tyr		
	515					520					525						
Arg	Ser	Met	Leu	Asp	Gln	Ile	Lys	Ser	Leu	Gly	Tyr	Asn	Thr	Ile	Arg		
	530				535					540							
Leu	Pro	Tyr	Ser	Asp	Asp	Ile	Leu	Lys	Pro	Gly	Thr	Met	Pro	Asn	Ser		
545					550					555					560		
Ile	Asn	Phe	Tyr	Gln	Met	Asn	Gln	Asp	Leu	Gln	Gly	Leu	Thr	Ser	Leu		
			565					570						575			
Gln	Val	Met	Asp	Lys	Ile	Val	Ala	Tyr	Ala	Gly	Gln	Ile	Gly	Leu	Arg		
		580						585				590					
Ile	Ile	Leu	Asp	Arg	His	Arg	Pro	Asp	Cys	Ser	Gly	Gln	Ser	Ala	Leu		
	595					600					605						
Trp	Tyr	Thr	Ser	Ser	Val	Ser	Glu	Ala	Thr	Trp	Ile	Ser	Asp	Leu	Gln		
	610				615					620							
Ala	Leu	Ala	Gln	Arg	Tyr	Lys	Gly	Asn	Pro	Thr	Val	Val	Gly	Phe	Asp		
625					630					635					640		
Leu	His	Asn	Glu	Pro	His	Asp	Pro	Ala	Cys	Trp	Gly	Cys	Gly	Asp	Pro		
			645					650						655			
Ser	Ile	Asp	Trp	Arg	Leu	Ala	Ala	Glu	Arg	Ala	Gly	Asn	Ala	Val	Leu		
		660						665				670					
Ser	Val	Asn	Pro	Asn	Leu	Leu	Ile	Phe	Val	Glu	Gly	Val	Gln	Ser	Tyr		
	675						680					685					
Asn	Gly	Asp	Ser	Tyr	Trp	Trp	Gly	Gly	Asn	Leu	Gln	Gly	Ala	Gly	Gln		
	690				695					700							
Tyr	Pro	Val	Val	Leu	Asn	Val	Pro	Asn	Arg	Leu	Val	Tyr	Ser	Ala	His		
705					710					715					720		
Asp	Tyr	Ala	Thr	Ser	Val	Tyr	Pro	Gln	Thr	Trp	Phe	Ser	Asp	Pro	Thr		
			725					730						735			
Phe	Pro	Asn	Asn	Met	Pro	Gly	Ile	Trp	Asn	Lys	Asn	Trp	Gly	Tyr	Leu		
		740						745				750					
Phe	Asn	Gln	Asn	Ile	Ala	Pro	Val	Trp	Leu	Gly	Glu	Phe	Gly	Thr	Thr		
	755						760					765					
Leu	Gln	Ser	Thr	Thr	Asp	Gln	Thr	Trp	Leu	Lys	Thr	Leu	Val	Gln	Tyr		
	770				775						780						
Leu	Arg	Pro	Thr	Ala	Gln	Tyr	Gly	Ala	Asp	Ser	Phe	Gln	Trp	Thr	Phe		
785					790					795					800		
Trp	Ser	Trp	Asn	Pro	Asp	Ser	Gly	Asp	Thr	Gly	Gly	Ile	Leu	Lys	Asp		
			805					810						815			
Asp	Trp	Gln	Thr	Val	Asp	Thr	Val	Lys	Asp	Gly	Tyr	Leu	Ala	Pro	Ile		
		820						825				830					
Lys	Ser	Ser	Ile	Phe	Asp	Pro	Val	Gly									
	835						840										

<211> 10239  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> pTrex plasmid

<400> 19

aagcttaact	agtacttctc	gagctctgta	catgtccggt	cgcgacgtac	gcgtatcgat	60
ggcgccagct	gcaggcggcc	gcctgcagcc	acttgacgtc	ccgtggaatt	ctcacgggtga	120
atgtaggcct	tttgtagggg	aggaattgtc	actcaagcac	ccccaacctc	cattacgcct	180
cccccataga	gttcccaatc	agtgagtcac	ggcactgttc	tcaaatagat	tggggagaag	240
ttgacttccg	cccagagctg	aaggctcgac	aaccgcacga	tataggggtc	gcaacggcaa	300
aaaagcacgt	ggctcaccga	aaagcaagat	gtttgcgac	taacatccag	gaacctggat	360
acatccatca	tcacgcacga	ccactttgat	ctgctggtaa	actcgtattc	gccctaaacc	420
gaagtgcagt	ggtaaactca	cacgtggggc	cctttcggta	tactgcgtgt	gtcttctcta	480
gggtgccattc	ttttcccttc	ctctagtgtt	gaattgtttg	tggtggagtc	cgagctgtaa	540
ctacctctga	atctctggag	aatggtggac	taacgactac	cgtgcacctg	catcatgtat	600
ataatagtga	tcctgagaag	gggggtttgg	agcaatgtgg	gactttgatg	gtcatcaaac	660
aaagaacgaa	gacgcctctt	ttgcaaagtt	ttgtttcggc	tacggtgaag	aactggatac	720
ttgtttgtgc	ttctgtgtat	ttttgtggca	acaagaggcc	agagacaatc	tattcaaaca	780
ccaagcttgc	tcttttgagc	tacaagaacc	tgtgggggat	atatctagag	ttgtgaagtc	840
ggtaatcccg	ctgtatagta	atacagagtc	catctaaata	ctccgaagct	gctgcgaacc	900
cggagaatcg	agatgtgctg	gaaagcttct	agcgagcggc	taaattagca	tgaaaggcta	960
tgagaaattc	tgagacgggc	ttgttgaatc	atggcggttc	attcttcgac	aagcaaagcg	1020
ttccgctcga	gtagcaggca	ctcattcccg	aaaaaactcg	gagattccta	agtagcgatg	1080
gaaccggaat	aatataatag	gcaatacatt	gagttgcctc	gacggttgca	atgcaggggg	1140
actgagcttg	gacataactg	ttccgtacct	cacctcttct	caacctttgg	cgtttccctg	1200
attcagcgta	cccgtacaag	tcgtaatcac	tattaaccca	gactgaccgg	acgtgttttg	1260
cccttcattt	ggagaaataa	tgtcattgct	atgtgttaatt	tgctgtgctg	accgactggg	1320
gctgttcgaa	gcccgaatgt	aggattgtta	tccgaactct	gctcgtagag	gcatgttgtg	1380
aatctgtgtc	gggcaggaca	cgctcgaag	gttcacggca	agggaaacca	ccgatagcag	1440
tgtctagtag	caacctgtaa	agccgcaatg	cagcatcact	ggaaaataca	aaccaatggc	1500
taaaagtaca	taagttaatg	cctaaagaag	tcataataca	gcggctaata	attgtacaat	1560
caagtggcta	aacgtaccgt	aatttgccaa	cggcttgttg	ggttgcagaa	gcaacggcaa	1620
agccccactt	ccccacgttt	gtttcttcac	tcagtcacat	ctcagctggg	gatcccccaa	1680
ttgggtcgct	tgttttgttc	gggtgaagtga	aagaagacag	aggtaagaat	gtctgactcg	1740
gagcgttttg	catacaacca	agggcagtg	tggaagacag	tgaaatgttg	acattcaagg	1800
agtatttagc	cagggatgct	tgagtgtatc	gtgtaaggag	gtttgtctgc	cgatacgacg	1860
aatactgtat	agtcacttct	gatgaagtgg	tccatattga	aatgtaagtc	ggcactgaac	1920
aggcaaaaaga	ttgagttgaa	actgcctaag	atctcggggc	ctcgggcctt	cggccttttg	1980
gtgtacatgt	ttgtgctccg	ggcaaagtga	aagtgtggta	ggatcgaaca	cactgctgcc	2040
tttaccaagc	agctgagggg	atgtgatagg	caaagtgtca	ggggccactg	catggtttcg	2100
aatagaaaaga	gaagcttagc	caagaacaat	agccgataaa	gatagcctca	ttaaacggaa	2160
tgagctagta	ggcaaagtca	gcgaatgtgt	atatataaag	gttcgaggtc	cgtgcctccc	2220
tcagtctctc	cccatctact	catcaactca	gactctccag	gagacttgta	caccatcttt	2280
tgaggcacag	aaacccaata	gtcaaccgag	gactgcgcac	catgtatcgg	aagttggccg	2340
tcattctcggc	cttcttgggc	acagctcggt	ctcagtcggc	ctgcactctc	caatcggaga	2400
ctcaccgcc	tctgacatgg	cagaaatgct	cgtctgggtg	cacttgcaact	caacagacag	2460
gctccgtggg	catcgacgcc	aactggcgct	ggactcacgc	tacgaacagc	agcacgaact	2520
gctacgatgg	caacacttgg	agctcgaccc	tatgtcctga	caacgagacc	tgcgcgaaag	2580
actgctgtct	ggacgggtgc	gcctacgcgt	ccacgtacgg	agttaccacg	agcggtaaca	2640
gcctctccat	tggctttgtc	acccagtcgt	cgcagaagaa	cgttggcgct	cgcctttacc	2700
ttatggcgag	cgacacgacc	taccaggaat	tcaccctgct	tggcaacgag	ttctcttttcg	2760
atgttgatgt	ttcgagctg	ccgtaagtga	cttaccatca	acccctgacg	tatcttcttg	2820
tgggctccca	gctgactggc	caatttaagg	tgcggttgta	acggagctct	ctacttcgtg	2880
tccatggacg	cggatgggtg	cgtgagcaag	tatcccacca	acaccgctgg	cgccaagtac	2940

ggcacggggt	actgtgacag	ccagtgtccc	cgcgatctga	agttcatcaa	tggccaggcc	3000
aacgttgagg	gctgggagcc	gtcatccaac	aacgcaaaca	cgggcattgg	aggacacgga	3060
agctgctgct	ctgagatgga	tatctgggag	gccaaactcca	tctccgaggc	tcttaccccc	3120
cacccttgca	cgactgtcgg	ccaggagatc	tgcgagggtg	atgggtgcgg	cggaaacttac	3180
tccgataaca	gatatggcgg	cacttgcgat	cccgatggct	gcgactggaa	cccataccgc	3240
ctgggcaaca	ccagcttcta	cggccctggc	tcaagcttta	ccctcgatac	caccaagaaa	3300
ttgaccgttg	tcacccagtt	cgagacgtcg	ggtgccatca	accgatacta	tgtccagaat	3360
ggcgtcactt	tccagcagcc	caacgccgag	cttggtagtt	actctggcaa	cgagctcaac	3420
gatgattact	gcacagctga	ggaggcagaa	ttcggcggat	cctctttctc	agacaagggc	3480
ggcctgactc	agttcaagaa	ggctacctct	ggcggcatgg	ttctggtcat	gagtctgtgg	3540
gatgatgtga	gtttgatgga	caaacatgcg	cgttgacaaa	gagtcaagca	gctgactgag	3600
atgttacagt	actacgccaa	catgctgtgg	ctggactcca	cctaccgcag	aaacgagacc	3660
tcctccacag	ccggtgccgt	gcgcggaagc	tgctccacca	gctccggtgt	ccctgctcag	3720
gtcgaatctc	agtctcccaa	cgccaaggctc	accttctcca	acatcaagtt	cggacccatt	3780
ggcagcaccg	gcaaccctag	cggcggcaac	cctcccggcg	gaaacccgcc	tggcaccacc	3840
accacccgcc	gcccagccac	taccactgga	agctctcccg	gacctactag	taagcggata	3900
aggcgcgcgg	cgcgccagct	ccgtgcgaaa	gcctgacgca	ccggtagatt	cttggtgagc	3960
ccgtatcatg	acggcggcgg	gagctacatg	gccccgggtg	atttatTTTT	tttgtatcta	4020
cttctgaccc	ttttcaaata	tacggtcaac	tcattcttca	ctggagatgc	ggcctgcttg	4080
gtattgcgat	gttgtcagct	tggcaaattg	tggctttcga	aaacacaaaa	cgattcctta	4140
gtagccatgc	attttaagat	aacggaatag	aagaaagagg	aaattaaaaa	aaaaaaaaaa	4200
acaaacatcc	cgttcataac	ccgtagaatc	gccgctcttc	gtgtatccca	gtaccagttt	4260
attttgaata	gctcgcccgc	tggagagcat	cctgaatgca	agtaacaacc	gtagaggctg	4320
acacggcgagg	tgttgctagg	gagcgtcgtg	ttctacaagg	ccagacgtct	tcgcggttga	4380
tatatatgta	tgtttgactg	caggctgctc	agcgacgaca	gtcaagttcg	ccctcgctgc	4440
ttgtgcaata	atcgcagtg	ggaagccaca	ccgtgactcc	catctttcag	taaagctctg	4500
ttggtgttta	tcagcaatac	acgtaattta	aactcgtag	catggggctg	atagcttaat	4560
taccgtttac	cagtgccgcg	gttctgcagc	tttccttgge	ccgtaaaatt	cggcgaagcc	4620
agccaatcac	cagctaggca	ccagctaaac	cctataatta	gtctcttata	aacaccatcc	4680
gctcccccg	gatcaatgag	gagaatgagg	gggatgcggg	gctaaacaag	cctacataac	4740
cctcatgcc	actcccagtt	tacactcgtc	gagccaacat	cctgactata	agctaacaca	4800
gaatgcctca	atcctgggaa	gaactggccg	ctgataagcg	cgcccgctc	gcaaaaacca	4860
tcctgatga	atggaaaagtc	cagacgctgc	ctgcggaaga	cagcgttatt	gatttcccaa	4920
agaaatcggg	gatcctttca	gaggccgaac	tgaagatcac	agaggcctcc	gctgcagatc	4980
ttgtgtccaa	gctggcggcc	ggagagttag	cctcggtgga	agttacgcta	gcattctgta	5040
aacgggcagc	aatcgcccag	cagttagtag	ggteccctct	acctctcagg	gagatgtaac	5100
aacgccacct	tatgggacta	tcaagctgac	gctggcttct	gtgcagacaa	actgcgcca	5160
cgagttcttc	cctgacgcgg	ctctcgcgca	ggcaagggaa	ctcgatgaat	actacgcaaa	5220
gcacaagaga	cccgttggtc	cactccatgg	cctccccatc	tctctcaaag	accagcttcg	5280
agtcaaggta	caccgttgcc	cctaagtcgt	tagatgtccc	tttttgtcag	ctaacatatg	5340
ccaccagggc	tacgaaacat	caatgggcta	catctcatgg	ctaaacaagt	acgacgaagg	5400
ggactcgggt	ctgacaacca	tgctccgcaa	agccggtgcc	gtcttctacg	tcaagacctc	5460
tgtcccgcag	accctgatgg	tctgcgagac	agtcaacaac	atcatcgggc	gcaccgtcaa	5520
cccacgcaac	aagaactggt	cgtgcggcgg	cagttctggt	ggtgagggtg	cgatcggttg	5580
gattcgtggt	ggcgtcatcg	gtgtaggaac	ggatatcggt	ggctcgattc	gagtgcgggc	5640
cgcgttcaac	ttcctgtacg	gtctaaggcc	gagtcatggg	cggctgccgt	atgcaaagat	5700
ggcgaacagc	atggagggtc	aggagacggg	gcacagcggt	gtcggggccga	ttacgcactc	5760
tgttgagggt	gagtccttcg	cctcttcctt	cttttcctgc	tctataccag	gcctccactg	5820
tcctcctttc	ttgcttttta	tactatatac	gagaccggca	gtcactgatg	aagtatgtta	5880
gacctccgcc	ttttcaccaa	atccgtcctc	ggtcaggagc	catggaaata	cgactccaag	5940
gtcatcccca	tgccctggcg	ccagtccgag	tcggacatta	ttgcctccaa	gatcaagaac	6000
ggcgggctca	atatcggtca	ctacaacttc	gacggcaatg	tccttccaca	ccctcctatc	6060
ctgcgcggcg	tggaaaccac	cgtcgccgca	ctcgccaaag	ccggtcacac	cgtgaccccg	6120
tggacgcat	acaagcacga	tttcggccac	gatctcatct	cccatatcta	cgcggctgac	6180
ggcagcggcg	acgtaatgcg	cgatatcagt	gcattccggcg	agccggcgat	tccaaatatc	6240
aaagacctac	tgaaccgcaa	catcaaagct	gttaacatga	acgagctctg	ggacacgcat	6300
ctccagaagt	ggaattacca	gatggagtag	cttgagaaat	ggcgggaggc	tgaagaaaag	6360

gccgggaagg	aactggacgc	catcatcgcg	ccgattacgc	ctaccgctgc	ggtagggcat	6420
gaccagttcc	ggtactatgg	gtatgcctct	gtgatcaacc	tgctggattt	cacgagcgtg	6480
gttgttccgg	ttacctttgc	ggataagaac	atcgataaga	agaatgagag	tttcaaggcg	6540
gttagtgagc	ttgatgccct	cgtgcaggaa	gagtatgata	cggaggcgta	ccatggggca	6600
ccggttgacg	tgcaggttat	cggacggaga	ctcagtgaag	agaggacgtt	ggcgattgca	6660
gaggaagtgg	ggaagtgtgt	gggaaatgtg	gtgactccat	agctaataag	tgtcagatag	6720
caatttgcac	aagaaatcaa	taccagcaac	tgtaaataag	cgctgaagtg	accatgccat	6780
gctacgaaa	agcagaaaaa	aacctgccgt	agaaccgaag	agatatgaca	cgcttccatc	6840
tctcaaagga	agaatccctt	cagggttgcg	tttccagtct	agacacgtat	aacggcacia	6900
gtgtctctca	ccaaatgggt	tatatctcaa	atgtgatcta	aggatggaaa	gcccagaatc	6960
taggcctatt	aatattccgg	agtatacgta	gccgggtaac	gttaacaacc	ggtacctcta	7020
gaactatagc	tagcatgcgc	aaattttaag	cgctgatata	gatcgcgcg	agatccatat	7080
atagggcccg	ggttataatt	acctcaggtc	gacgtcccat	ggccattcga	attcgtaatc	7140
atgggtcatag	ctgtttcctg	tgtgaaattg	ttatccgctc	acaattccac	acaacatacg	7200
agccggaagc	ataaagtgtg	aagcctgggg	tgcctaatag	gtgagctaac	tcacattaat	7260
tgcgttgcgc	tactgccccg	ctttccagtc	gggaaacctg	tcgtgccagc	tgcattaatg	7320
aatcggccaa	cgcgcgggga	gaggcggttt	gcgtattggg	cgctcttccg	cttctctcgt	7380
cactgactcg	ctgcgctcgg	tcgttcgggt	gcggcgagcg	gtatcagctc	actcaaaggc	7440
ggtaatacgg	ttatccacag	aatcagggga	taacgcagga	aagaacatgt	gagcaaaagg	7500
ccagcaaaag	gccaggaacc	gtaaaaaggc	cgcgttgctg	gcgtttttcc	ataggctccg	7560
ccccctgac	gagcatcaca	aaaatcgacg	ctcaagtcag	aggtggcgaa	acccgacagg	7620
actataaaga	taccaggcgt	ttccccctgg	aagctccctc	gtgcgctctc	ctgttccgac	7680
cctgccgctt	accggatacc	tgtccgcctt	tctcccttcg	ggaagcgtgg	cgctttctca	7740
tagctcacgc	tgtaggtatc	tcagttcggg	gtaggtcgtt	cgctccaagc	tgggctgtgt	7800
gcacgaaccc	cccgttcagc	ccgaccgctg	cgccttatcc	ggtaactatc	gtcttgagtc	7860
caacccggta	agacacgact	tatcgccact	ggcagcagcc	actggttaaca	ggattagcag	7920
agcgaggtat	gtaggcggtg	ctacagagtt	cttgaagtgg	tggcctaact	acggctacac	7980
tagaagaaca	gtattttggt	tctgcgctct	gctgaagcca	gttaccttcg	gaaaaagagt	8040
tggtagctct	tgatccggga	aacaaaccac	cgtgggtagc	ggtgggtttt	ttggttgcaa	8100
gcagcagatt	acgcgcagaa	aaaaaggatc	tcaagaagat	cctttgatct	tttctacggg	8160
gtctgacgct	cagtggaacg	aaaactcacg	ttaagggtat	ttggtcatga	gattatcaaa	8220
aaggatcttc	acctagatcc	ttttaaatca	aaaatgaagt	tttaaatcaa	tctaaagtat	8280
atatgagtaa	acttgggtctg	acagttacca	atgcttaatc	agtgaggcac	ctatctcagc	8340
gatctgtcta	tttcgttcat	ccatagttgc	ctgactcccc	gtcgtgtaga	taactacgat	8400
acgggagggc	ttaccatctg	gccccagtcg	tgcaatgata	ccgcgagacc	cacgctcacc	8460
ggctccagat	ttatcagcaa	taaaccagcc	agccggaagg	gccgagcgca	gaagtgggtc	8520
tgcaacttta	tccgcctcca	tccagtctat	taattgttgc	cgggaagcta	gagtaagtag	8580
ttcgccagtt	aatagtttgc	gcaacgttgt	tgccattgct	acaggcatcg	tggtgtcacg	8640
ctcgctgctt	ggtatggctt	cattcagctc	cggttcccaa	cgatcaaggc	gagttacatg	8700
atcccccatg	ttgtgcaaaa	aagcgggttag	ctccttcggt	cctccgatcg	ttgtcagaag	8760
taagttggcc	gcagtgttat	cactcatggt	tatggcagca	ctgcataatt	ctcttactgt	8820
catgccatcc	gtaagatgct	tttctgtgac	tggtgagtag	tcaaccaagt	cattctgaga	8880
atagtgtatg	cggcgaccga	gttgctcttg	cccggcgctc	atacgggata	ataccgcgcc	8940
acatagcaga	acttttaaaag	tgctcatcat	tggaaaacgt	tcttcggggc	gaaaactctc	9000
aaggatctta	ccgctgttga	gatccagttc	gatgtaaccc	actcgtgcac	ccaactgatc	9060
ttcagcatct	tttactttca	ccagcgtttc	tgggtgagca	aaaacaggaa	ggcaaaatgc	9120
cgcaaaaaag	ggaataaggg	cgacacggaa	atgttgaata	ctcactactc	tcctttttca	9180
atattattga	agcattttat	agggttattg	tctcatgagc	ggatacatat	ttgaatgtat	9240
ttagaaaaat	aaacaaatag	gggttccgcg	cacatttccc	cgaaaagtgc	cacctgacgt	9300
ctaagaaaac	attattatca	tgacattaac	ctataaaaaat	aggcgatatca	cgaggccctt	9360
tcgtctcgcg	cgtttcgggtg	atgacggtga	aaacctctga	cacatgcagc	tcccgagac	9420
ggtcacagct	tgtctgtaag	cggatgccgg	gagcagacaa	gcccgtcagg	gcgcgtcagc	9480
gggtgttggc	gggtgtcggg	gctggcttaa	ctatgcggca	tcagagcaga	ttgtactgag	9540
agtgacccat	aaaattgtaa	acgttaatat	tttggttaaaa	ttcgcggttaa	atttttgtta	9600
aatcagctca	ttttttaacc	aataggccga	aatcggaaca	atcccttata	aatcaaaaga	9660
atagcccag	ataggggtga	gtgttggtcc	agtttggaac	aagagtccac	tattaaagaa	9720
cgtggactcc	aacgtcaaa	ggcgaaaaac	cgtctatcag	ggcgatggcc	cactacgtga	9780

accatcaccc	aaatcaagtt	ttttggggtc	gaggtgccgt	aaagcactaa	atcggaaccc	9840
taaaggaggc	ccccgattta	gagcttgacg	gggaaagccg	gcgaacgtgg	cgagaaagga	9900
aggggaagaaa	gcgaaaggag	cgggcgctag	ggcgctggca	agtgtagcgg	tcacgctgcg	9960
cgtaaccacc	acacccgccg	cgcttaatgc	gccgctacag	ggcgcgctact	atggttgctt	10020
tgacgtatgc	ggtgtgaaat	accgcacaga	tgcgtaagga	gaaaataccg	catcaggcgc	10080
cattcgccat	tcaggctgcg	caactgttgg	gaagggcgat	cgggtgcgggc	ctcttcgcta	10140
ttacgccagc	tggcgaaagg	gggatgtgct	gcaaggcgat	taagttgggt	aacgccaggg	10200
ttttcccagt	cacgacgttg	taaaacgacg	gccagtgcc			10239

<210> 20  
 <211> 12  
 <212> DNA  
 <213> Trichoderma reesei

<400> 20	
actagtaagc gg	12

<210> 21  
 <211> 41  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> primer

<400> 21	
gcttatacta gtaagcgcgc gggcggcggc tattggcaca c	41

<210> 22  
 <211> 39  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> primer

<400> 22	
gccttatggcg cgccttagac aggatcgaaa atcgacgac	39

<210> 23  
 <211> 43  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> primer

<400> 23	
ctaagagaaa cgacccgtac atccagcggc tcctcagcat gta	43

<210> 24  
 <211> 38  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> primer

<400> 24  
 ttaccccgat gggaagagca tgccaaaatc ggcgttcg 38

<210> 25  
 <211> 43  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> primer

<400> 25  
 ctaagagagc gacgactcag ccgtacacct ggagcaacgt ggc 43

<210> 26  
 <211> 38  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> primer

<400> 26  
 ttacgatccg gacggcgcac caccaatgtc cccgtata 38

<210> 27  
 <211> 42  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> primer

<400> 27  
 gcttatacta gtaagcgcg cggctgctcg gtggactaca cg 42

<210> 28  
 <211> 37  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> primer

<400> 28  
 gcttatggcg cgccttacag aggcgggtag gcgttgg 37

<210> 29  
 <211> 42  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> primer

<400> 29

gcttatacta gtaagcgcg cggcttcacc gccacagtca cc

42

<210> 30

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 30

gcttatggcg cgcctcagga ctggagcttg ctccgc

36n

21